

Definition, Convention, and Simultaneity: Malament's Result and Its Alleged Refutation by Sarkar and Stachel

Robert Rynasiewicz

Department of Philosophy, Johns Hopkins University

The question whether distant simultaneity (relativized to an inertial frame) has a factual or a conventional status in special relativity has long been disputed and remains in contention even today. At one point it appeared that Malament (1977) had settled the issue by proving that the only non-trivial equivalence relation definable from (temporally symmetric) causal connectability is the standard simultaneity relation. Recently, however, Sarkar and Stachel (1999) claim to have identified a suspect assumption in the proof by defining a non-standard simultaneity relation from causal connectability. I contend that their critique is based on a misunderstanding of the criteria for the definability of a relation, a misunderstanding that Malament's original treatment helped to foster. There are in fact a variety of notions of definability that can be brought to bear. They all, however, require a condition that suffices to secure Malament's result. The non-standard relation Sarkar and Stachel claim to be definable is not so definable, and, I argue, their proposal to modify the notion of "causal definability" is misguided. Finally, I address the relevance of Malament's result to the thesis of conventionalism.

1. Introduction. In his classic paper setting out the special theory of relativity, Einstein (1905) observed that, although the time of an event can be given by the simultaneous reading of a clock in its immediate neighborhood, the times of occurrence of events at different locations cannot be compared without some further arrangement [Festsetzung] (278-9). One has, at best, "time" given by the readings of a clock at the one location and a "time" given by those of a clock at the other. A common "time" can be defined by fixing *by definition* that the "time" it takes for light to travel from the one to the other is equal to the "time" required for the reverse. From this follows the standard synchronization condition that the two clocks synchronize if, upon sending a light ray from one to the other and reflecting it back to the first, the reading on the clock at the site of reflection exactly splits the difference between the readings on the first clock at departure and at return (279).

Although the rationale for postulating *by definition* the isotropy of the speed of light was lost on not a few early commentators and textbook writers (see in particular Silberstein (1914, 93–99), Schlick (1920, 14–17) and Cunningham (1921, 28–31)), Reichenbach (1924, 10) eventually assimilated Einstein’s isotropy definition to the status of a convention in the sense in which Helmholtz and Poincaré had held that geometry is conventional. Grünbaum (1973, 342–69) understood the basis for the Einstein definition to lie the “fact” that the relation of causal connectability between spacetime events is insufficient for uniquely establishing a strict temporal ordering of events, even from the point of view of a fixed inertial frame of reference.

Malament (1977) directly challenges this alleged “fact.” According to Malament, Grünbaum commits himself to the following pair of assertions (293):

- (1) The relation of simultaneity relative to an inertial observer is not uniquely definable in terms of the relation κ of causal connectability.
- (2) Temporal relations are non-conventional if and only if they are so definable.

Malament observes that A. A. Robb (1914) had in effect already shown that the relation of orthogonality on the vector space R^4 with Minkowski inner product is “explicitly, first order definable in terms of κ ” (296), where the ordered pair (p, q) is in the relation κ just in case the inner product of p with q is non-negative.¹ Hence, one can define two events e_1, e_2 to be simultaneous relative to inertial observer O if and only if for some (in fact, any) pair r, s of distinct events on the world-line of O , the vector $e_2 - e_1$ is orthogonal to $r - s$. This is equivalent to the standard Einstein synchronization condition above. Malament goes on to show that the only other non-trivial² equivalence relation between events definable from κ and O is the universal relation. Hence, assuming that the relation of simultaneity for an inertial observer is a non-trivial equivalence relation smaller than the universal relation, it follows that the standard simultaneity relation is uniquely definable, contra (1) above. By (2), then, it is non-conventional.

Recently, Sarkar and Stachel (1999) have challenged the validity of Malament’s result. They claim to have found another (in fact, two other) non-trivial equivalence relations definable for O from κ . From the fact that these are ruled out by Malament’s proof they infer that Malament imposes “an unwarranted physical assumption: that any simultaneity relation must remain invariant under temporal reflections.” (208), and argue that this assumption should be relaxed. Spirtes (1981, 171–186) has also claimed that if the restriction of invariance under temporal reflections is lifted, then non-standard simultaneity relations are definable, in fact uncountably many.

The purpose of this note is to explain in what ways Sarkar and Stachel’s critique of Malament is misguided. The non-standard simultaneity relation they claim to be able to define for O from κ is in fact not so definable. Furthermore,

¹This assumes the sign convention $(+, -, -, -)$.

²In the sense that at least some point on O is equivalent to at least one point not on O .

Malament’s result does not rely on “physical assumptions” either warranted or unwarranted. Unfortunately, the issue is somewhat confused by the fact that Malament gave an incorrect formulation of what it means for a relation to be implicitly definable. Sarkar and Stachel assume that this formulation has the status of a *proposal* and thus consider it open for criticism whether what they call Malament’s “explication” (Sarkar and Stachel, 212) is “physically reasonable” (215). However, just what relations are definable from what is a purely formal, mathematical question and has nothing to do with issues of physical adequacy. There are in fact several, distinct formal notions of definability available from mathematical logic, and it is less than clear specifically which of these Malament had in mind. However, no matter which is adopted, his result remains unaffected, since it relies on a condition strictly weaker than the one he gives, and this weaker condition is a logical consequence of each of these notions of definability. In conclusion, I shall offer a few remarks of the bearing of Malament’s result on the issue of the conventionality of simultaneity.

2. Sarkar and Stachel on the Definability of Light Lobes. One of the results due to Robb (1914) is that the relation of *lightlike relatedness* λ between events in Minkowski spacetime is definable from κ alone.³ Quite clearly then, any relation that is definable using λ in addition is definable from κ alone. In particular, Sarkar and Stachel claim that the relation “ p and q lie on the same half null cone with vertex e ” is just so definable in virtue of the following defining condition:

- (i) p and q both lie on the light cone with vertex e , and
- (ii) either p and q are not causally connectable or else there exists a signal connecting them which does not pass through the space-like “elsewhere” of e or through e itself.

³The relation λ on R^4 is the set of ordered pairs $\{(p, q) \mid (p - q) \cdot (p - q) = 0\}$. The first order formula

$$p\kappa q \wedge \exists r(r \neq p \wedge r \neq q \wedge \forall s(s\kappa p \wedge \kappa q \rightarrow s\kappa r))$$

defines λ in the structure (R^4, κ) . For those expert in these matters who might be put off by the lack of complete precision in this section and the next, I offer the following remarks. (For those not expert, the technicalities will explained in section 4.) Above I have used the notion of definability in a structure rather than that of explicit definability in a theory. Malament’s locution “explicitly first order definable in terms of κ (300) is ambiguous between this and the notion of explicit (first order) definability in a theory. As will be seen later, this formula also explicitly defines λ in T , where T can be characterized equivalently as either the deductive closure of Robb’s axioms or the set of all sentences of the predicate language containing κ alone that are true in (R^4, κ) . Note as well that, following Malament, I have systematically failed to distinguish between ‘ κ ’ as a name for the predicate of this language and ‘ κ ’ as a symbol for the relation assigned to that predicate as its interpretation in the structure (R^4, κ) . Also, in order not to disrupt the flow of the discussion, I will, for the time being, continue to use in the body of the text such locutions as “definable from κ ” despite the ambiguity as to which notion of definability is intended.

The first task is to check whether indeed these conditions can be expressed rigorously using only κ and/or λ , something Sarkar and Stachel do not concern themselves with. The first conjunct (i) is easily rendered by the formula:

$$(i') \quad p\lambda e \wedge q\lambda e.$$

The second conjunct (ii) requires a bit of finesse. The trouble lies with the invocation of the notion of a causal signal, i.e., a continuous curve whose tangent is never spacelike. *Epistemologically* (and maybe even ontologically), it certainly makes sense to regard this notion as *more* primitive than the relation of causal connectability. However, in the game that's at the table, in which causal connectability is the *sole* primitive, it must, rigorously speaking, be reconstructed from that sole primitive. Actually, for the purposes at hand, it would suffice to find a way of expressing when three events are respectively optically, or inertially collinear, and when so, which of the three is between the other two. Most of this task can be handled in the fashion of Robb via definitions of *optical line*, *parallel optical lines*, *acceleration plane*, *inertial line*, etc. However, since Robb took the notion of *causally after* to be his primitive, some additional cunning is needed, particularly in handling relations of betweenness. The details are edifying, but hardly to the point for present purposes.

The second, and main task is to assess what Sarkar and Stachel think follows from this. What bearing does this have on Malament's result? Sarkar and Stachel argue as follows. Temporarily fix the parameter e and let $\Lambda(e)$ be the full light cone centered at e .

It follows that there are two such half null cones, with only e in common, that is, $\Lambda_-(e) \cup \Lambda_+(e) = \Lambda(e)$ and $\Lambda_-(e) \cap \Lambda_+(e) = \{e\}$. Note that this distinction, once made on the null cone at any point of spacetime, can be consistently carried by parallel transport to the null cone at any other point of spacetime.

Thus λ can be used to define two additional simultaneity relations, $Back_{O,e}$ and $Forw_{O,e}$, relative to a time-like world line O : two events are simultaneous relative to an event e on O if and only if they lie on $\Lambda_-(e)$ [$\Lambda_+(e)$]. Clearly neither relation is "vacuous" in Malament's (297) sense: e is simultaneous to events not on O and not all events are simultaneous. Each is also trivially an equivalence relation. Most importantly, each of the relations can be defined from λ alone, since $\Lambda_-(e)$ [$\Lambda_+(e)$] can be so defined. (Note, that since they are definable from λ , they are also definable from κ .)

Clearly, something is amiss with Malament's theorem. A correct mathematical result seems to be contradicted by patently good counterexamples.

What may be patent to them may not be so for everyone. Again, we need to check for rigor. To begin with, there is an obvious slip, though one that is

easily correctable. Fix the parameters O and e in the definition of $Back_{O,e}$. The resulting equivalence relation has but two equivalence classes, $\Lambda_-(e)$ and its complement. Hence, any event not on the past light cone of e is “ $Back_{O,e}$ ”-simultaneous to any other event not on the back light cone of e . Clearly this is not what is intended. The proposed simultaneity criterion should yield a one-parameter family of equivalence relations, one for each choice of world line O . Having fixed O , the events e on O then generate the equivalence classes $[e]$ of the relation defined. So, the term ‘ e ’ should not appear as an additional parameter in the definiens, but at best as a variable quantifier over in the definiendum.

The relation Sarkar and Stachel have in mind is in fact given by:

$$Sim_{\bar{O}}(p, q) \text{ iff } \exists e(e \in O \wedge p, q \in \Lambda_-(e)).$$

In other words, p and q are back-light-cone simultaneous for world-line O just in case they lie on the back light cone of some event e on O (and similarly for the forward directed counterpart relation $Sim_{\bar{O}}^+$).

These two relations, however, are *not* counterexamples; they are *not* definable from λ and/or κ alone. The reason is that, although the relation of lying on the same half null cone may be definable, and indeed it can be shown that there exist exactly two half cones, neither one of these cones is itself definable. It is an elementary fallacy to think that just because two things satisfying a certain condition are given, one thereby succeeds in defining either one of them merely by introducing a pair of labels. In the event that this is not completely obvious, it may help to examine some parallel examples.

3. Two Analogues. Consider from group theory Klein’s *Viergruppe*, the four-element group V with the following multiplication table, where e is the identity element.⁴

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

The automorphisms of V are given by all permutations of a, b, c . Although V contains exactly three elements distinct from the identity element, a fact obviously first-order expressible in the language of group theory, none of these elements is individually definable. (The language can be taken to include, besides the usual logical symbols, an individual constant symbol for the identity

⁴Up to isomorphism, there is only one four-element group, a fact potentially disguised by some of its realizations. For example, Klein’s Viergruppe is also the symmetry group of a line segment in the Euclidean plane, as well as the product group $C_2 \times C_2$, where C_n is the cyclic group of order n .

element and a function symbol for group multiplication.) To drive home the general point, we can construct a tighter analogy with the light cone case (with fixed vertex e) as follows. The entire group V will play the role of $\Lambda(e)$. Next define a relation xRy (which you are invited to read as ‘ x lies on the same “lobe” as y ’) by

$$xRy \text{ iff } ((x \circ y = x) \vee (x \circ y = y)).$$

The relation R generates the three “lobes” $\{e, a\}$, $\{e, b\}$, $\{e, c\}$, which as a triplet you are invited to think of as the analogue of the pair of cones with vertex e . Note, again, that there exist exactly three “lobes” sharing the identity element is first-order expressible.⁵ Now, in order for the lobe, say $\{e, a\}$ to be definable, there must exist a group-theoretic equation in one free variable which is solved by e and a alone. However, from the fact that the automorphisms of V are the permutations of a, b, c , it follows that if a is a solution to an equation, then so are b and c . Hence, no individual lobe is definable. I hope this helps to make it more compelling why it is that simply introducing labels gets one no closer to definitions.

A second example, which matches the light lobes case in cardinality, relies only on the topology of the real line. Let betweenness be given as the only primitive relation on R^1 and pick some distinguished point of R^1 , for the sake of definiteness the base e of the natural logarithm, as the only other primitive. Now define the relation $S^e xy$ by

$$S^e xy \text{ iff } \neg eBxy,$$

where $zBxy$ is to be read “ z lies between x and y .” Hence $S^e xy$ just in case x and y lie on the same ray with end point e . Think of this binary relation as the analogue of “ p and q lie on the same half null cone with vertex e ”. Obviously, there are exactly two rays of the real line with end point e a fact expressible even without recourse to a many sorted language or higher order quantification.⁶ Think of these two rays as counterparts of the two lobes of the light cone with vertex at event e . Of course, we might introduce labels, say RAY_- and RAY_+ “merely to distinguish them” (Sarkar and Stachel, 213), without claiming to know which denotes which ray or in anyway thereby introducing an orientation on the real line. But again, this in no way establishes that each individual ray is definable from the betweenness relation (and e) alone. What is needed for each is a predicate expression (involving only e and B) in one free variable satisfiable

⁵Transcribe into first order notion the following: there exist at least three elements, none on the same lobe as another, such that any other element is on the same lobe as one of the three.

⁶Consider the first-order sentence

$$\exists x \exists y (x \neq y \wedge x \neq e \wedge y \neq e \wedge S^e x e \wedge S^e y e \wedge \neg S^e x y \wedge \forall z (z \neq x \wedge z \wedge y \wedge z \neq e \rightarrow S^e z x \vee S^e z y)).$$

A first order sentence of the same form then expresses the fact that there are exactly two half light cones at e by substituting the analogous predicate and relativizing the quantifiers to the set of events lightlike related to e .

by all and only points on the ray to be defined. As will shortly become evident, it is provable by symmetry arguments that no such expressions exist.

In presenting these examples, I hope to have tapped intuitions about the conditions under which a given relation or object is definable. But I do not want the case to rest solely with that. There is a standard literature on definability that backs those intuitions. Before turning to this, it may be instructive to consider just what does transpire in the familiar form of reasoning which proceeds: “We have shown there are exactly two things which are such-and-such. Call them c_1 and c_2 respectively. It now follows that . . .”

Let $\varphi(x)$ be a predicate expression in one free variable x , and suppose it has been established that there are exactly two things satisfying $\varphi(x)$. What is the effect of introducing arbitrary tags c_1, c_2 in subsequent reasoning? The upshot of the move is to make available the following assumptions: $\varphi(c_1), \varphi(c_2), c_1 \neq c_2$, and finally $\forall x(\varphi(x) \rightarrow (x = c_1 \vee x = c_2))$. If one then succeeds in deriving a sentence σ not containing either c_1 or c_2 , then one knows that σ can be proven, in a probably much more obtuse fashion, without the introduction of c_1 and c_2 and their associated assumptions. Those of you stuck with teaching logic term after term will recognize this as nothing other than a generalized instance of the rule of inference called Existential Instantiation, according to which, having shown $\exists x\varphi(x)$, one says “let us introduce a name c for an arbitrarily selected individual satisfying $\varphi(x)$ ” and proceeds to employ the assumption $\varphi(c)$ in subsequent reasoning. No one, of course, thinks that one has thereby succeeded in *defining* one of the entities that happens to satisfy $\varphi(x)$.

I turn now to the standard literature on definability.

4. Notions of Definability. So far I have been somewhat sloppy with regard to certain technical distinctions, among them, just what notion of definability is in play. The preceding discussion of h , however, is most naturally construed as addressing the notion of *definability in a structure*.

Let \mathfrak{A} be a structure for a first-order language \mathcal{L} , R an n -ary relation on the domain of discourse of \mathfrak{A} , and s an assignment of elements of the domain of discourse of \mathfrak{A} to the individual variables of \mathcal{L} . Then R is said to be definable in \mathfrak{A} just in case the following: there is a formula φ of \mathcal{L} whose free individual variables are among x_1, \dots, x_n such that for any s , \mathfrak{A} satisfies φ with s if and only if the n -tuple $\langle s(x_1), \dots, s(x_n) \rangle$ is a member of R (Enderton 1972, 88).

In order to illustrate how this applies in the above examples, we need to be slightly more precise in our notation. For the group-theoretic example, the non-logical vocabulary contains the two-place function symbol \circ for group composition and an individual constant symbol e for the group identity.⁷ The Viergruppe \mathfrak{V} , as a structure for this language, consists of the set $V = \{e^{\mathfrak{V}}, a, b, c\}$ together with the binary operation $\circ^{\mathfrak{V}}$ specified by the multiplication table.

⁷Although the group identity is definable in any model of group theory, I include e as a primitive of the language for expository purposes.

(The horrendous notation is necessary to distinguish the identity *symbol* from the identity *element* and the function *symbol* from the symbol's *interpretation*.) Since, technically speaking, the relation R is not a predicate symbol, but a subset of $V \times V$, the definition of R in this example should read:

$$\langle a_1, a_2 \rangle \in R \text{ iff } \mathfrak{V} \models ((x \circ y = x) \vee (x \circ y = y)) [a_1, a_2],$$

where $\mathfrak{V} \models \varphi(x, y)[a_1, a_2]$ means that \mathfrak{V} satisfies $\varphi(x, y)$ with s for any assignment s of elements of v to variables such that $s(x) = a_1$ and $s(y) = a_2$.

The question as to whether a given subset (e.g., a “lobe”) of the Viergruppe is first order definable in the group is now just the question as to whether there is a defining formula with one free variable. The following lemma can be used to establish that certain subsets (or relations in general) are *not* definable in a given structure.

First Automorphism Lemma. The (in general n -ary) relation R is definable in \mathfrak{A} only if R is preserved under all automorphisms of \mathfrak{A} .

It immediately follows that none of the three “lobes” of the Viergruppe is definable in the group since they are not invariant under group automorphisms. Similarly with the two rays from the second example. Reflection through the point e preserves betweenness on the real line, and hence is an automorphism of that structure. But neither ray is preserved under reflections.

One might ask about higher-order definability in these examples. Fortunately, it is not necessary to go into the details of the conditions for second or higher order logic. The above automorphism lemma holds under generalization to higher orders.

There are two other standard notions of definability, both variants of the idea of *definability in a given theory*, where a theory is understood to be a deductively closed set of sentences. Here, the distinction between a predicate symbol and its interpretation as a relation in a given model of the theory is crucial since what is involved is really the definition of the intension of a predicate (or function) symbol. By the intension of an n -ary predicate in a theory, we mean a function that assigns to each model of the theory an n -ary relation on the domain of discourse of the model. When it is said that a given relation is definable in a theory, what is meant is that the intension of an associated predicate symbol is definable. It is also taken for granted that the associated predicate is part of the vocabulary in which the theory is formulated. However, so much nit-pickiness eventually becomes cumbersome and, despite the potential for types confusions, it is simply said that: an n -ary relation P is *explicitly* definable in theory T if and only if in the language of T there exists a formula $\theta(x_1, \dots, x_n)$, with variables x_1, \dots, x_n free but in which P does not occur, such that the sentence

$$\forall x_1 \dots \forall x_n (Px_1, \dots, x_n \leftrightarrow \theta(x_1, \dots, x_n))$$

is a theorem of T . In order to relate this to that of to the previous notion of definability in a structure, we adopt the following notational convention. If \mathfrak{A} is

a structure for the language of T minus the n -ary predicate P and R is an n -ary relation on the domain of discourse of \mathfrak{A} , then (\mathfrak{A}, R) is a structure for the full language of T in which R is the extension of P . The following is not difficult to prove: if P is explicitly definable in T , then the relation R is definable in \mathfrak{A} for any model (\mathfrak{A}, R) of T . In fact, the open formula $\theta(x_1, \dots, x_n)$ that explicitly defines P in T defines R in \mathfrak{A} . An immediate consequence is a corollary to the above automorphism lemma.

Second Automorphism Lemma. P is explicitly definable in theory T only if for each model (\mathfrak{A}, R) of T the extension R of P is preserved under automorphisms of \mathfrak{A} .

The other standard notion of definability in a theory is that of implicit definability.⁸ Using the above notation, P is said to be *implicitly* definable in T just in case if (\mathfrak{A}, R) and (\mathfrak{A}, R') are any two models of T , then $R = R'$. A result due to Beth is that for first order logic P is implicitly definable in T if and only if P is explicitly definable in T . For higher order logics only the “if” direction of the biconditional holds generally. Nonetheless, a direct and general connection between implicit definability and model automorphism is available, so general in fact that it applies directly to classes of structures without the intermediary of deductively closed sets of sentences if we replace the models of T with elements of a class K of structures in the characterization of implicit definability.

Third Automorphism Lemma. If P is implicitly definable in K and K is closed under isomorphism, then the extension R of P in any member (\mathfrak{A}, R) of K is preserved under automorphisms of \mathfrak{A} .

The proof is rather immediate. Let (\mathfrak{A}, R) be a member of K and ϕ an automorphism of \mathfrak{A} . Letting ϕ act on (\mathfrak{A}, R) yields the isomorphic structure $(\phi*\mathfrak{A}, \phi*R)$, which is also a member of K if K is closed under isomorphism. However, since ϕ is an automorphism of \mathfrak{A} , $(\phi*\mathfrak{A}, \phi*R) = (\mathfrak{A}, \phi*R)$. Thus, if P is implicitly definable in K , $(\mathfrak{A}, \phi*R) = (\mathfrak{A}, R)$.

5. Definable Relations in Minkowski Spacetime. The punch line is now this. No matter what notion of definability is in question, the preservation of a relation under automorphisms of the structure or structures in question is a necessary condition for the definability of the relation. In the case of Minkowski spacetime, these automorphisms include temporal reflections and thus render futile any attempt to define an individual half cone in the absence of a temporal orientation. *A fortiori* the proposed simultaneity relation $Sim_{\bar{O}}$ (or $Sim_{\bar{O}}^{\dagger}$) is not definable “in terms of κ (and O) alone” whether this is parsed as meaning

⁸The notion of implicit definability dealt with here should not be confused with that in the writings of Schlick and Reichenbach on the foundations of mathematics inspired by Hilbert’s axiomatization of geometry.

definable in the structure (R^4, κ, O) , or as implicitly definable in the class of structures isomorphic to $(R^4, \kappa, O, Sim_{\bar{O}})$, or as explicitly definable in the set of first or any higher order sentences true in $(R^4, \kappa, O, Sim_{\bar{O}})$.

So, it might be asked, why overcome this by providing Minkowski spacetime with a temporal orientation at the outset? After all, Robb took a (temporally) *asymmetric* causal connectability relation as his primitive, and CP violation in weak interactions certainly suggests that fundamental laws will not be time-reversal invariant. This is what Sarkar and Stachel do in effect by proposing an alternative “definition” of definability. It is unclear, though, what they think is the upshot of this move.

Part of the difficulty stems from a bit of terminology that slips in innocently towards the beginning of their paper, but is quickly perverted into an expression that is either pleonastic or pernicious. In order to talk about those mappings of R^4 to itself that preserve the relation of causal connectability, Malament introduces the term of art “causal automorphism” (297). Sarkar and Stachel describe Malament as “explicit[ing] the notion of ‘causally definable’ ” (212). As far as I have been able to determine, the latter expression never occurs in Malament’s paper. Nonetheless, it could serve as reasonable shorthand for “definable from κ .” However, within the span of a paragraph Sarkar and Stachel introduce further the locution “causally definable from κ and O ,” which appears time and again in what follows without further elucidation. As a substitute for “causal automorphism” in the sense just explained, this is pleonistic — “causally definable from O ” suffices. If some other reading is intended, it is pernicious — for then there is no precedent whatsoever for what it might mean.

No doubt Malament can be accused of having invited confusion by misstating the condition for definability in Minkowski spacetime. He writes that a relation is “implicitly definable from κ ” IF and only if it is preserved under all causal automorphisms (297).⁹ This led Spirtes (1981), solely on the mistaken assumption that preservation under causal automorphisms is a *sufficient* for definability, to the conclusion that in a temporally oriented Minkowski spacetime there are uncountably many non-standard simultaneity relations definable from κ and the temporal orientation.¹⁰ Malament’s *proof*, it should be stressed, requires only the “only if” direction. Inclusion of the “if” direction yields a criterion far too promiscuous to of general service. That is, no further interesting sense of “definability” is to be had with the criterion: R is definable in structure \mathfrak{A} iff R is invariant under all automorphisms of \mathfrak{A} . To see this, consider the standard model of arithmetic. It has no non-trivial automorphisms. Hence, it follows from the criterion that *any* arithmetic relation is definable, a consequence that certainly begs the question in what sense one would at all want to speak of “definability.”

To sum up, Sarkar and Stachel’s proposal to redefine the notion “relation

⁹Emphasis mine.

¹⁰As far as I am aware, the backward and forward light cone relations are the only two additional equivalence relations definable in a temporally oriented Minkowski spacetime.

R is causally definable from κ and O ” is doubly misguided. First, we have no idea what qualification the term “causally” is supposed to add. Second, as an instance of the suggestion to formulate some alternative notion of definability, it falls prey to the sorts of worries just raised in connection with arithmetic.

6. Conclusion. The validity of Malament’s result remains unaffected by the considerations and objections raised by Stachel and Sarkar. Despite what they say, the *only* assumption Malament requires to refute assertion (1), i.e., to establish that the standard (frame-relative) simultaneity relation is uniquely definable from causal connectability, is that the relation be a non-trivial equivalence relation strictly smaller than the universal relation.

As to the *relevance* of Malament’s result for the issue of the conventionality of distant simultaneity in special relativity, there remain two obvious questions. One is whether distant simultaneity must be an equivalence relation. That it need not be has been suggested by Reichenbach (1924) and Redhead (1983). The other is whether the conventionalist is committed to assertion (2) from Malament’s paper, i.e., that temporal relations are non-conventional if and only if they are uniquely definable from causal connectability.

Concerning the first of these, it is not entirely beside the point that in his 1905 paper, Einstein proposes as a definition that the time required for light to travel from point A to point B equals the time required for it to travel from B to A . It can be shown that the principle of relativity and the light postulate are consistent with alternative definitions. However, if it is further required that space is homogenous and isotropic, then the Einstein definition is a necessary consequence of the fundamental postulates (Rynasiewicz, notes). *Within* the special theory of relativity, there are really no alternatives to the standard simultaneity criterion.

As for the second, it should be noted that in general there are perhaps a half dozen distinct forms of conventionalism (Rynasiewicz, 1999). So, whether or not the conventionalist is committed to (2) does not admit a short answer. (Whether Grünbaum commits himself to (2), I leave for him to say.) It is worth pointing out, though, that neither Reichenbach nor Einstein would have accepted (2). The question for them was not what we are at liberty to say about distant simultaneity after the theory of special relativity is given, but rather to what degree one is constrained prior to the articulation of the theory.¹¹ Einstein introduces the velocity definition within the context of the myth of the “stationary” frame prior to any application of the theory of relativity. It is meant to provide a criterion acceptable to aether-theorist and relativist alike. (With modest elaboration it is also acceptable to the emission-theorist, but it is unlikely Einstein was concerned to address this possibility in the 1905 paper even

¹¹See (Stein 1991, 152–155) for a similar diagnosis.

though he had explored that option the preceding year or two. See Rynasiewicz, 2000.) If indeed this is correct, then most of the discussions since simply miss the mark.

REFERENCES

- Chang, C. C. and H. J. Keisler (1973), *Model Theory*. Amsterdam: North Holland.
- Cunningham, E. (1921), *Relativity, the Electron Theory, and Gravitation*, 2nd edition. London: Longmans, Green and Co.
- Einstein, Albert (1905), “Zur Elektrodynamik bewegter Körper”, *Annalen der Physik* 17: 891–921.
- Enderton, Herbert B. (1972), *A Mathematical Introduction to Logic*. San Diego: Academic Press
- Grünbaum, Adolf, (1973), *Philosophical Problems of Space and Time*, 2nd edition. Dordrecht: Reidel.
- Malamet, David (1977), “Causal Theories of Time and the Conventionality of Simultaneity”, *Noûs* 11: 293–300.
- Reichenbach, Hans (1924), *Axiomatik der relativistischen Raum-Zeit-Lehre*, Braunschweig: Viewig Verlag. Translated into English by Maria Reichenbach as, *Axiomatization of the Theory of Relativity*. Berkeley: University of California Press, 1969.
- Rob, A. A. (1914), *A Theory of Time and Space*. Cambridge: Cambridge University Press.
- Rynasiewicz, Robert (1999). “Varieties of Conventionality,” *Volume of Abstracts*, International Union of History and Philosophy of Science, and 11th International Congress of Logic, Methodology and Philosophy of Science. Edited by Jacek Cachro and Katarzyna Kijania-Placek. Cracow: Kopiorama. Page 329.
- Rynasiewicz, Robert (2000). “The Construction of the Special Theory: Some Queries and Considerations.” In *Einstein: The Formative Years, 1879–1909*. Edited by Don Howard and John Stachel. Boston: Birkhäuser. Pages 159–201.
- Rynasiewicz, Robert (notes). “The Houdini Aether Theory.”
- Sarkar, Sahotra and John Stachel (1999), “Did Malament Prove the Non-Conventionality of Simultaneity in the Special Theory of Relativity?”, *Philosophy of Science* 66: 208–220.
- Schlick, Moritz (1920), *Raum und Zeit in der gegenwärtigen Physik*, 3rd edition. Berlin: Julius Springer.
- Silberstein, L. (1914), *The Theory of Relativity*. London: Macmillan.
- Spirtes, Peter L. (1981), *Conventionalism and the Philosophy of Henri Poincaré*. Ph. D. Dissertation, Department of the History and Philosophy of Science, University of Pittsburgh.
- Stein, Howard (1991), “On Relativity Theory and Openness of the Future”, *Philosophy of Science* 58: 147–167.